



EP3260: Machine Learning Over Networks
Computer Assignment 3
Due Date: March 7, 2023

Computer Assignment 3 - Training a neural network

Consider optimization problem

$$\underset{\mathbf{W}_1, \mathbf{W}_2, \mathbf{w}_3}{\text{minimize}} \quad \frac{1}{N} \sum_{i \in [N]} \|\mathbf{w}_3 \mathbf{s}(\mathbf{W}_2 \mathbf{s}(\mathbf{W}_1 \mathbf{x}_i) - \mathbf{y}_i)\|_2^2,$$

where $\mathbf{s}(\mathbf{x}) = 1/(1 + \exp(-\mathbf{x}))$. You may add your choice of regularizer. Using the “Individual household electric power consumption” and “Greenhouse Gas Observing Network” datasets, address the following questions:

- (a) Try to solve this optimization task with proper choices of size of decision variables (matrix \mathbf{W}_1 , matrix \mathbf{W}_2 , and vector \mathbf{w}_3) using GD, perturbed GD, SGD, SVRG, and block coordinate descent. For the SGD method, you may use the mini-batch version.
- (b) Compare these solvers in terms complexity of hyper-parameter tuning, convergence time, convergence rate (in terms of # outer-loop iterations), and memory requirement